ATTACHMENT A

Eur. J. Biochem. **268,** 5961–5969 (2001) © FEBS 2001

Characterization of two DNA polymerases from the hyperthermophilic euryarchaeon *Pyrococcus abyssi*

Yannick Gueguen[1], Jean-luc Rolland[1], Odile Lecompte[2], Philippe Azam[1], Gisèle Le Romancer[1], Didier Flament[1], Jean-Paul Raffin[1] and Jacques Dietrich[1]

[1]*IFREMER, Centre de Brest, DRV-VP-LBMH, Plouzané, France;* [2]*IGBMC, Laboratoire de Biologie Structurale, Illkirch, France*

The complete genome sequence of the hyperthermophilic archaeon *Pyrococcus abyssi* revealed the presence of a family B DNA polymerase (Pol I) and a family D DNA polymerase (Pol II). To extend our knowledge about euryarchaeal DNA polymerases, we cloned the genes encoding these two enzymes and expressed them in *Escherichia coli*. The DNA polymerases (Pol I and Pol II) were purified to homogeneity and characterized. Pol I had a molecular mass of ≈ 90 kDa, as estimated by SDS/PAGE. The optimum pH and Mg^{2+} concentration of Pol I were 8.5–9.0 and 3 mM, respectively. Pol II is composed of two subunits that are encoded by two genes arranged in tandem on the *P. abyssi* genome. We cloned these genes and purified the Pol II DNA polymerase from an *E. coli* strain coexpressing the cloned genes. The optimum pH and Mg^{2+} concentration of Pol II were 6.5 and 15–20 mM, respectively. Both *P. abyssi* Pol I and Pol II have associated $3'\rightarrow5'$ exonuclease activity although the exonuclease motifs usually found in DNA polymerases are absent in the archaeal family D DNA polymerase sequences. Sequence analysis has revealed that the small subunit of family D DNA polymerase and the Mre11 nucleases belong to the calcineurin-like phosphoesterase superfamily and that residues involved in catalysis and metal coordination in the Mre11 nuclease three-dimensional structure are strictly conserved in both families. One hypothesis is that the phosphoesterase domain of the small subunit is responsible for the $3'\rightarrow5'$ exonuclease activity of family D DNA polymerase. These results increase our understanding of euryarchaeal DNA polymerases and are of importance to push forward the complete understanding of the DNA replication in *P. abyssi*.

Keywords: *Pyrococcus abyssi*; archaea; DNA polymerase; DNA replication; exonuclease activity.

DNA polymerases play a leading role in the replication and maintenance of the genome and are central to the accurate transmission of genetic information from generation to generation. While our knowledge about DNA replication in eukarya and bacteria is quite advanced [1], limited information is available on the replication mechanism in archaea, the third major domain of life [2]. Recently, comparative genomics revealed that most archaeal proteins involved in DNA replication, transcription and translation are similar to those in eukarya, although the cellular appearance and organization of archaea are more similar to bacteria. Recent investigations [3,4] revealed that within the archaeota, euryarchaeota and crenarchaeota, the two major subdomains differ in their DNA replication mechanisms. The analysis of genome sequences indicated that many euryarchaea [5–7] possess a single family B (α-like) DNA-polymerase. In addition, a new heterodimeric DNA-polymerase (family D DNA polymerase) [8], with no significant homology to eukaryal or bacterial DNA polymerases, has been detected in the corresponding euryarchaeal genomes. No homologues of this heterodimeric DNA polymerase has so far been detected in the crenarchaeota kingdom and the genome sequencing of *Aeropyrum pernix* [9] and *Sulfolobus solfataricus* [10] suggests the absence of family D DNA polymerase homologues in crenarchaeotic cells. In contrast, the existence of two family B DNA polymerases in *A. pernix* [11] and *Pyrodictium occultum* [12] and of three family B DNA polymerases in *S. solfataricus* [13] indicates that several B-type DNA polymerases exist in the crenarchaeotal genomes. These findings confirm that the DNA replication mechanism of the euryarchaeal and crenarchaeal subdomains of archaea differs, and therefore opens the discussion of the evolution of DNA polymerases, a group of indispensable proteins that are central to the replication process. However, for a more precise interpretation of the evolutionary relationship between archaea and eukarya, additional genome sequences from crenarchaea require investigation.

P. abyssi is an anaerobic hyperthermophilic archaeon that belongs to the Thermococcales order within the euryarchaea subdomain. It was isolated from hydrothermal vents at a depth of 2000 meters in the south-west Pacific Ocean and grows optimally around 100 °C [14]. The complete genome sequence of *P. abyssi* revealed the presence of two DNA polymerases: a family B DNA polymerase (Pol I) and a family D DNA polymerase (Pol II). To increase our knowledge about euryarchaeal DNA polymerases, we cloned the genes for these two DNA polymerases, expressed them in

Correspondence to Y. Gueguen, IFREMER-CNRS, UMR 219, Defense et Resistance chez les Invertebres Marins (DRIM), 2 Place E. Bataillon, CP 80, F-34095 Montpellier cedex 5, France.
Fax: + 33 4 67 14 46 22, Tel.: + 33 4 67 14 47 07,
E-mail: ygueguen@ifremer.fr

Abbreviations: Pol I, DNA polymerase I; Pol II, DNA polymerase II; PCNA, proliferating cell nuclear antigen; RPA, replication protein A.

(Received 11 July 2001, revised 20 September 2001, accepted 26 September 2001)

Escherichia coli and characterized the two recombinant DNA polymerases.

MATERIALS AND METHODS

Organisms and growth conditions

P. abyssi (strain Orsay) was used in this study. The complete genome sequence of *P. abyssi* was determined at Genoscope (Evry, France). Sequences and annotations are available at http://www.genoscope.cns.fr/Pab/. The entire nucleotide sequence of *P. abyssi* was submitted to EMBL database under accession numbers CNSPAX01 to CNSPAX06. *E. coli* HMS174 (DE3), which harbors pLysS, was used as a host strain for overexpressing the cloned genes in the recombinant plasmids pPOLI, pPOLB and pPOLC. *E. coli* DH5α was also used as a host strain for the subcloning step. *E. coli* was grown in Luria–Bertani medium in a rotary shaker at 37 °C with ampicillin and/or kanamycin (final concentration of 100 $\mu g \cdot mL^{-1}$ or 30 $\mu g \cdot mL^{-1}$, respectively).

Cloning of *P. abyssi* Pol I (*polI*) and Pol II genes (*polB, polC*)

Based on the sequence of *pol1* (accession number P77916), primers were designed to amplify the DNA polymerase I gene (*pol1*) from *P. abyssi* by the PCR on a DNA Thermal Cycler (Stratagene). The two primers (with *Nde*I and *Bam*HI restriction sites in bold) were as follows: POLI-1 (sense), 5′-CAGATTGGGTGGGG**CATATG**ATAATCG ATGC-3′; and POLI-2 (antisense), 5′-CCCGA**GGATCC** TAGAACTTAAGCCATGCTCC-3′. The heterodimeric DNA Pol II from *P. abyssi* is composed of two subunits (PolB and PolC) whose genes (*polB* and *polC*) are arranged in tandem on the *P. abyssi* genome (Fig. 1). Based on the sequences of *polB* and *polC* (accession numbers F75199 and E75199, respectively), primers were designed to amplify separately the two DNA polymerase II genes. The two sets

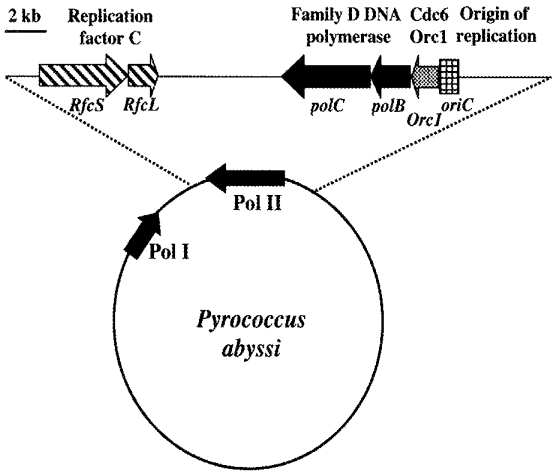

Fig. 1. Localization of the genes corresponding to Pol I and Pol II in the genome of *P. abyssi*. Genes located in the vicinity of Pol II that encode proteins that participate in DNA replication are shown. These include the origin of replication, the two subunits of the replication factor C and the Cdc6/Orc1 protein that is needed for the initiation of DNA replication.

of primers (with *Nde*I and *Bam*HI or *Nde*I and *Sal*I restriction sites in bold) were as follows: POLB1 (sense), 5′-CAAAGGAGGTTGCT**CATATG**GATGAATTGGTT AAGG-3′; POLB2 (antisense), 5′-TTCCTTTGGA**GGAT CC**ATCAACACCACCCGCTG-3′; POLC1 (sense), 5′-AG CGGGTGGTG**CATATG**GAGCTTCCAAAGG-3′; POLC2 (antisense), 5′-TCGATGAGTACTAAG**GTCGAC**TTAGT AGATTTCACG-3′. In addition to the DNA template from *P. abyssi* and the primers, the 50-μL reaction mixture contained 10 nmol dNTPs, *Pfu* DNA polymerase buffer and 2 U *Pfu* DNA polymerase (Promega) and was subjected to 20 cycles of amplification (30 s at 94 °C, 30 s at 50 °C and 3 min at 72 °C). A PCR product of the expected size for *polI* was digested with *Nde*I and *Bam*HI and cloned into pET-26b + (Novagen, Inc.), resulting in pPOLI, and transformed into *E. coli* DH5α according to standard procedures [15]. The *polC* gene, encoding the large subunit of the Pol II, was shown to contain an intein coding region, which is an intervening sequence spliced out as a protein and not as a mRNA [22]. To prevent a possible toxic effect of the intein on *E. coli*, the gene encoding intein was deleted by splicing by overlap extension PCR [16] with the following primers: POLC3: 5′-GAGGAGAAACTGTGATGGAGAT GAAGACGCTG-3′ and POLC4: 5′-CTCCATCACAGTTT CTCCTCTTCGCAGCGTGG-3′, resulting in the *polC2* gene. The two fragments containing *polB and polC2* were digested with *Nde*I and *Bam*HI or *Nde*I and *Sal*I, respectively, and cloned into pARHS [17] or pET-26b+ (Novagen), respectively, resulting in pPOLB and pPOLC, which were transformed into *E. coli* DH5α according to standard procedures [15].

Expression and purification of *P. abyssi* Pol I and Pol II

The recombinant plasmid pPOL1 containing *pol1* gene was used for transformation of *E. coli* HMS174 (DE3)pLysS. In the case of Pol II, the two recombinant plasmids, pPOLB and pPOLC, were used to cotransform *E. coli* HMS174 (DE3)pLysS. Overnight cultures of *E. coli* HMS174 (DE3)pLysS harboring pPOL1 and *E. coli* HMS174 (DE3)pLysS harboring pPOLB and pPOLC were diluted 1 : 20 and grown until D_{600} reached 0.6. The two cultures were induced with 1 mM of isopropyl thio-β-D-galactoside (IPTG) for 16 h. Cells were harvested by centrifugation. The Pol I culture was resuspended in 50 mM phosphate buffer (pH 7.5 containing 1 mM EDTA, 10 mM 2-mercaptoethanol and 200 $\mu g \cdot mL^{-1}$ lysozyme) and the Pol II culture was resuspended in 50 mM Tris/HCl buffer (pH 8 containing 10% glycerol and 10 mM 2-mercaptoethanol). The cells were disrupted by sonication using a Vibracell sonifier (375 W, 40% amplitude). Cell debris were removed by centrifugation (10 000 ***g*** for 10 min). The resulting supernatants were heated for 30 min at 70 °C for Pol I and 30 min at 80 °C for Pol II and the precipitated proteins were removed by further centrifugation. The supernatants were subsequently dialyzed against buffer A (50 mM Tris/HCl, pH 8) for Pol I and buffer B (50 mM potassium phosphate, pH 6.5, 2 mM 2-mercaptoethanol, 10% glycerol) for Pol II. The dialysates were applied to a Resource-Q column (Pharmacia) that had been pre-equilibrated with the corresponding buffer (A for Pol I, B for Pol II) by use of an FPLC system (Pharmacia). Bound proteins were eluted by a linear gradient of NaCl (0–0.5 M

in buffer A or B). The active fractions were dialyzed against buffer A for Pol I and buffer B for Pol II and applied to a 5-mL heparin–Sepharose column (Hitrap Heparin, Pharmacia). The columns were developed with a linear gradient of 150 mM to 1 M NaCl in buffer C (50 mM Tris/HCl, pH 7.5, 1 mM dithiothreitol, 0.1 mM EDTA, 0.05% Triton X-100) for Pol I and a linear gradient of 150 mM to 650 mM in buffer B for Pol II.

Protein samples were analyzed by SDS/PAGE using the method of Laemmli [18]. Protein concentrations were determined by the method of Bradford, with BSA as the standard [19]. Protein bands were visualized by staining with Coomassie Brillant Blue R-250.

DNA polymerase assay

DNA polymerizing activity was assayed by measuring the incorporation of [*methyl*-^{3}H]dTMP into trichloroacetic acid insoluble material. The principle of the assay has been described previously [12,20]. In brief, the 20-μL assay mixture contained: (a) for Pol I, 50 mM Tris/HCl pH 8.8, 1 mM dithiothreitol, 10 mM KCl, 2 mM $MgCl_2$, 0.4 mg·mL^{-1} BSA, 20 μM [^{3}H]dTTP (1.5 Ci·mmol^{-1}), 0.42 μM of poly(dA)$_{265}$/oligo(dT)$_{17}$ (10 : 1), 200 μM (each) dATP, dGTP, dCTP and dTTP and 1 μL of the enzyme fraction; (b) for Pol II, 20 mM potassium phosphate (pH 6.5), 2 mM 2-mercaptoethanol, 10 mM $MgCl_2$, 20 μM [^{3}H]dTTP (1.5 Ci·mmol^{-1}), 0.42 μM of poly(dA)$_{265}$/oligo(dT)$_{17}$ (10 : 1), 200 μM (each) dATP, dGTP, dCTP and dTTP and 1 μL of the enzyme fraction. All reactions were incubated at 65 °C for 30 min. The amount of radioactivity incorporated into DNA strands was measured in a scintillation counter. One unit is defined as the amount of enzyme catalyzing the incorporation of 1 nmol of dTMP per min at 65 °C into polydA/oligodT.

Exonuclease activity assay

Associated exonuclease activities were assayed using previously described methods [12,21]. For the 3′→5′ exonuclease assay, pBluescript II SK(−) [pBS SK(−)] was digested by *Eco*RI, purified and the resultant linear fragment was labeled at the 3′ end by use of the Klenow enzyme in the presence of [^{3}H]dTTP. The 5′→3′ exonuclease activity was measured using a 446-bp *Eco*RI pBS SK(−) fragment labeled at the 5′ end with [γ-^{32}P]ATP and polynucleotide kinase. Two hundred and fifty nanogams of 3′-labeled DNA and 8 ng of 5′-labeled DNA were used for 3′→5′ exonuclease and 5′→3′ exonuclease assays, respectively. The 20-μL reaction mixture contained for (a) Pol I: 50 mM Tris/HCl pH 8.8, 1 mM dithiothreitol, 60 mM KCl, 2 mM $MgCl_2$, 0.4 mg·mL^{-1} BSA, 3′ or 5′-labeled DNA and 0.3 U of purified DNA polymerase; (b) Pol II: 20 mM potassium phosphate (pH 6.5), 2 mM 2-mercaptoethanol, 10 mM $MgCl_2$, 250 ng 3′- or 5′-labeled DNA and 0.3 U of purified DNA polymerase. The mixtures were incubated 65 °C for 90 min. The release of acid-soluble nucleotides was scored by adding 60 μL of trichloroacetic acid (10%) and 3 mg·mL^{-1} BSA to the 20-μL reaction. After incubating for 10 min on ice, the samples were centrifuged at 12 000 *g* for 5 min. The acid-soluble radioactivity in 60 μL of supernatant was quantified using Pico-Fluor 15 (Packard) scintillation fluid in a scintillation counter.

Optima pH, thermal stability, substrate specificities

The optimum pHs for the two DNA polymerases were determined by carrying out the standard assay at 65 °C, using (a) for Pol I: Tris/HCl (50 mM) and glycine/NaOH (50 mM) buffers for the pH range 7–10 and 8–11, respectively; (b) for Pol II: potassium phosphate (20 mM) and Tris/HCl (50 mM) buffers for the pH range 6–8 and 7–9.5, respectively. Thermostability was determined by use of purified enzymes (0.8 mg·mL^{-1} in buffer C for Pol I and 0.5 mg·mL^{-1} in buffer B for Pol II) and incubated at 70, 80, 90 and 100 °C for 2 h. Residual activity was determined by standard procedures at 65 °C. To compare the substrate specificities, of the two *P. abyssi* DNA polymerases, the standard assay was run with the following substrates: poly(dA)$_{265}$/oligo(dT)$_{17}$ (10 : 1) (0.42 μM, nucleotides concentration), primed M13 circular ssDNA (M13 DNA) (2.1 μM, nucleotide concentration) and activated calf thymus DNA (13 μg·mL^{-1}).

RESULTS AND DISCUSSION

Cloning of the two DNA polymerases genes from *P. abyssi*

The *pol1* gene, encoding a family B DNA polymerase, and the *polB* and *polC* genes, encoding a heterodimeric family D DNA polymerase [8] were identified among the 1765 ORFs in the complete genome sequence of *P. abyssi* (available at http://www.genoscope.fr/Pab/). The 2313-bp *pol1* gene is located at positions 1 695 183–1 697 495 on the *P. abyssi* chromosome. Its predicted product is a protein of 771 amino-acids, with a theoretical molecular mass of 84.8 kDa. In *P. abyssi*, *polB* (1857 bp) and *polC* (4365 bp) of the family D DNA polymerase, are arranged in tandem and located at positions 115 179–121 402 on the *P. abyssi* genome. The *PolC* gene, encoding the large subunit of Pol II, was shown to contain an intein which is an intervening sequence spliced out as a protein and not as a mRNA [22]. After the production of the precursor protein, the intein is excised from the protein. To prevent a possible toxic effect of the intein in *E. coli*, the

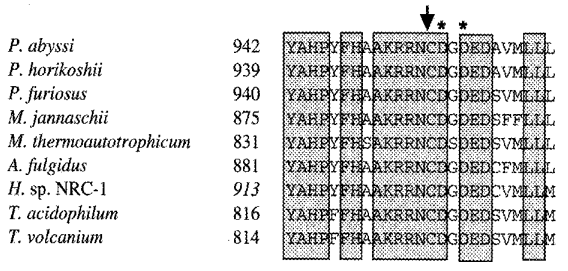

Fig. 2. Comparison of the conserved amino-acid sequence region in the active site of the family D DNA polymerases. The two proposed catalytic residues [27] are marked by an asterisk. The arrow indicates the position of the mini-intein insertion in *P. abyssi, P. horikoshii and Halobacterium* sp. NRC-1. Conserved identical residues are highlighted in grey. Sequences were deduced from the following accession numbers (SWISSPROT amino-acids sequence database): *P. horikoshii* (Q57861), *P. abyssi* (Q9V2F4), *P. furiosus* (P91409), *M. jannaschii* (*Q59024*), *M. thermoautotrophicum* (*O27579*), *A. fulgidus* (*O28552*), *H.* sp. NRC-1 (Q9HMX2), *T. acidophilum* (*Q9HM33*), *T. volcanium* (complete genome NC_002689, accession NP_110554).

555-bp gene encoding intein was deleted, resulting in *polC2*. Three of the nine family D DNA polymerase sequences from the database were shown to be inserted by a mini-intein: *P. abyssi* (185 amino-acids length), *P. horikoshii* (166 amino-acids in length) and *Halobacterium* sp. NRC-1 (195 amino-acids length) [27]. Interestingly, in these organisms, the mini-intein is inserted in the most conserved region of the family D DNA polymerase in close vicinity of the two catalytic aspartic residues (Fig. 2). Thus, the deduced product of *polB* and *polC2* genes are proteins of 619 and 1270 amino-and acids, with theoretical molecular masses of 68.1 and 139.7 kDa, respectively. The organization of the genes encoding the DNA polymerase D components is interesting (Fig. 1). The two genes responsible are arranged in tandem in the *P. abyssi* genome and the regions flanking *polB* and *polC* contain genes that encode proteins that are essential for DNA replication. These include a homologue of the initiator protein Orc1/Cdc6, the two subunits of the polymerase accessory complex replication factor C and the identified origin of replication (*oriC*) [23,24]. The organization of these genes is similar in the genus *Pyrococcus* (*P. furiosus*, *P. woesei and P. horikoshii*) [8]. Although the family D of DNA polymerases has only been detected in the euryarchaeal subdomain of archaea, the tandem organization of the genes encoding the two subunits of the heterodimeric polymerase seems to be specific to the Thermococcales: *Pyrococcus* (*P. furiosus, P. woesei* [8] and *P. abyssi*, present work) and *Thermococcus* (*T. aggregans*, K. Böhlke, Institute of Technical Microbiology, Hamburg, Germany, personal communication). The significance of these differences in gene arrangement is not known. However, we expect that the organization as an operon could be advantageous to regulate the level of the two DNA polymerase subunits in the cells.

Comparison of the amino-acid sequences of Pol I and Pol II with other DNA polymerases

The sequences of the two *P. abyssi* DNA polymerases were aligned with those of archaeal DNA polymerases available in public databases. The amino-acid sequence of Pol I contains the six conserved motifs shared by the family B DNA polymerases and the three motifs for $3'\rightarrow5'$ exonuclease activity [13]. The Pol I sequence from *P. abyssi* showed highest identities to the euryarchaeal DNA polymerases from the family B than to the crenarchaeal family B DNA polymerases (Table 1). These results are consistent with the report on the phylogenetic relationships of the archaeal family B DNA polymerases [13]. Comparison of the deduced primary sequences of *polB* and *polC* with the sequences of enzymes present in public databases indicated that they are highly similar to representatives of the family D DNA polymerases, which only contains heterodimeric DNA polymerases from euryarchaea. So far, only nine family D DNA polymerase sequences have been deposited in the genetic database. The two subunits of the heterodimeric polymerases are highly conserved within the representatives of the DNA polymerases family D, suggesting that these proteins play an important role in the organisms. Comparison of *P. abyssi* Pol II to members of family D revealed identities of between 79% (*P. horikoshii*) and 28% (*M. jannashii*) for the small subunits and between 91% (*P. horikoshii*) and 44% (*Thermoplasma acidophilum*) for the large subunit. In addition, as previously reported in other euryarchaea [25], the small subunit of *P. abyssi* has significant similarities to the small subunit of the eukaryotic pol δ and, to a lesser extent, to the small subunit of the eukaryotic pol α and pol ε. Furthermore, although the large subunit of the family D DNA polymerases exhibits no similarity to the catalytic domain of any known DNA polymerases from the database, amino-acid sequence analysis showed that it contains conserved motifs, including invariant carboxylates found in the palm domain of the polymerase superfamily [8,26]. However, site directed mutagenesis of this residue showed that they are not involved in the DNA polymerization activity of *P. horikoshii* family D DNA polymerase [27]. In the same study, two aspartate residues, located in a highly conserved region of the large subunit were proposed, based on site directed

Table 1. Comparison of the amino-acid sequences of archaeal family B DNA polymerases. Identity values were calculated using the GAP program (GCG Wisconsin Package) in pairwise comparison. Nomenclature of the polymerase (B1 and B3) is that of Edgell *et al.* [13].

	Amino-acid identity (%)										
Organism	*Pab*	*Pho*	*Pfu*	*Tli*	*Mja*	*Afu*	*Poc* B3	*Sso* B3	*Poc* B1	*Sso* B1	*Spo*
Euryarchaea											
P. abyssi	100	92	84	76	47	37	35	25	26	26	29
P. horikoshii		100	82	75	46	37	34	24	26	25	29
P. furiosus			100	74	47	37	35	25	25	27	28
T. litorallis				100	47	38	33	24	25	26	28
M. jannashii					100	36	32	24	23	26	25
A. fulgidus						100	43	29	24	25	26
Crenarchaea											
P. occultum B3							100	33	26	25	27
S. solfataricus B3								100	18	18	20
P. occultum B1									100	50	19
S. solfataricus B1										100	19
Eukarya											
S. pombe PolD											100

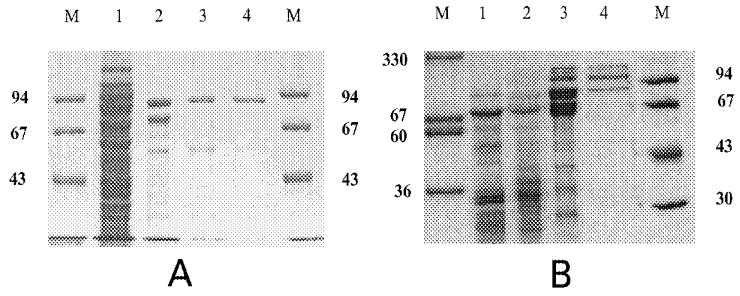

Fig. 3. SDS/PAGE of enzymatic fractions generated during the purification of Pol I (A) and Pol II (B). Lane 1, sonicated crude extract; lane 2, supernatant after heat treatment; lane 3, Resource-Q fraction, lane 4, heparin–Sepharose fraction. Lanes labeled 'M' contain size marker proteins (HMW and LMW calibration kits, Pharmacia). Molecular masses are indicated in the margins.

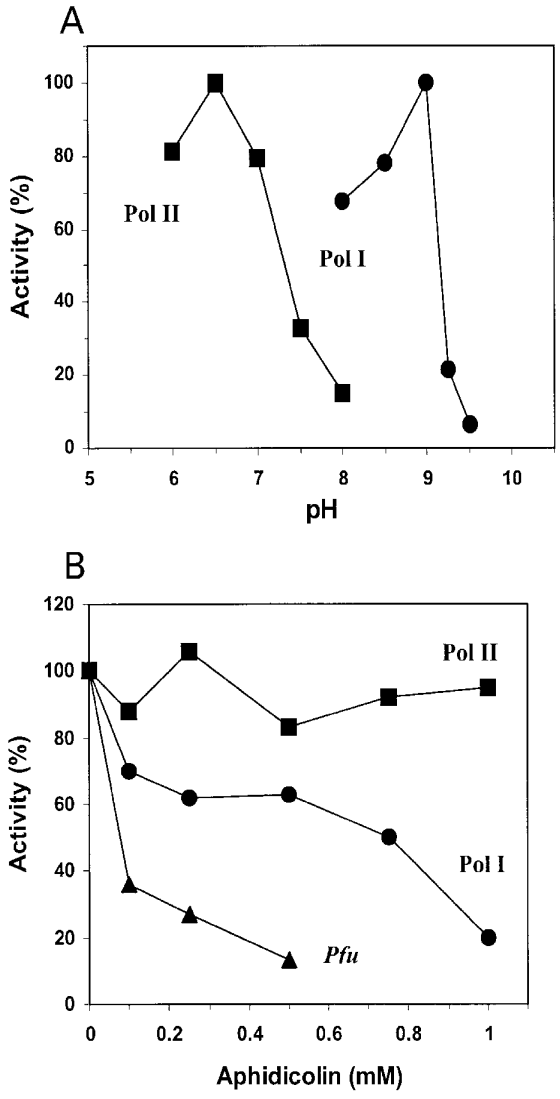

Fig. 4. Influence of pH (A) and aphidicolin (B) on the activities of *P. abyssi* Pol I and Pol II. (A) The standard assay was carried out at 65 °C for 30 min in 50 mM glycine-NaOH buffer for Pol I and in 20 mM potassium phosphate buffer for Pol II. One hundred percent activity corresponds to 80 U·mg^{-1} of protein for Pol I and 17 U·mg^{-1} of protein for Pol II. (B) The standard assay was carried out at 65 °C for 30 min with various concentrations of aphidicolin. One hundred percent activity corresponds to 74 U·mg^{-1} of protein for Pol I and 17 U·mg^{-1} of protein for Pol II. *Pfu*, *P. furiosus* DNA polymerase.

mutagenesis, to be the catalytic residues of the family D DNA polymerase (Fig. 2) [27].

Expression and purification of *P. abyssi* Pol I and Pol II

To determine the properties of the two *P. abyssi* DNA polymerases, the *polI*, *polB* and *polC2* genes were cloned into pET-26b+, pARHS and pET-26b+, respectively, under the control of the T7 promoter. The resulting plasmids were checked by DNA sequence analysis. The two DNA polymerases were produced in recombinant form in *E. coli* HMS174(DE3)pLysS cells after IPTG induction. Pol I and Pol II were purified to near-homogeneity by a three-step procedure: (a) incubation at a high temperature that denatures most of the *E. coli* proteins; (c) anion exchange; and (c) Hi-Trap-heparin column chromatography (data not shown). SDS/PAGE analysis of the homogeneous proteins revealed a single band of ≈90-kDa for Pol I (Fig. 3A) corresponding to the calculated molecular mass of the *polI* product and three bands of ≈140, 110 and 90 kDa for Pol II (Fig. 3B). The two subunits of the heterodimeric Pol II were also expressed separately in *E. coli* HMS174(DE3), pLysS (data not shown). The 90-kDa band corresponded to the small subunit of Pol II, although its predicted molecular mass was 68.1 kDa. Similar results were obtained for the small subunit of the heterodimeric DNA polymerase of *P. furiosus* [28] and *P. horikoshii* [27]. The N-terminal sequences of the 140- and 110-kDa bands were analyzed after electroblotting from a denaturating gel to a poly(vinylidene difluoride) membrane. These sequences were found to be identical to the N-terminal sequence of the large subunit of Pol II. The 140-kDa band is the same size as the calculated molecular mass of *polC2* product and the 110-kDa band probably corresponds to a degradation product of the 140-kDa large subunit of Pol II. So far, only three family D DNA polymerases have been characterized [27–29]. The different studies showed that purification of the *P. furiosus* and *P. horikoshii* family D DNA polymerases resulted in the rapid degradation of the small subunit for *P. furiosus* and the large subunit *P. horikoshii* during the different steps [27,28]. This suggests that family D DNA polymerases are highly unstable during the purification procedure. *P. abyssi* Pol II migrated as a 230-kDa band in native PAGE, which is equal to the sum of the molecular masses of the deduced amino-acid sequences of the two subunits. This suggests that, as for *P. furiosus*, the active DNA polymerase complex is a heterodimer made up of a 1 : 1 ratio of the two Pol II subunits [8,28]. On the other hand, based on gel filtration results, the structure of *P. horikoshii* recombinant family D DNA polymerase was shown to be different; the results suggested a

Table 2. Activities of DNA polymerases on various template primers. *Pfu*, *P. furiosus* DNA polymerase. *Taq*, *Taq* DNA polymerase.

	Relative activity			
Substrate	Pol I	Pol II	*Pfu*	*Taq*
Calf thymus-activated DNA	1.00	1.00	1.00	1.00
M13 ssDNA 30-mer primer	2.48	2.85	0.79	2.83
Poly(dA)-oligo(dT) (10 : 1)	7.45	1.92	3.13	2.18

Table 3. Comparison of the exonuclease activities associated with DNA polymerases. One unit of DNA polymerizing activity is defined as described in Materials and methods. *Pfu*, *P. furiosus* DNA polymerase. *Taq*, *Taq* DNA polymerase. ND, non detectable.

		Relative nucleolytic activity (%)	
DNA polymerase	DNA polymerizing activity (U)	3′→5′ nucleolytic	5′→3′ nucleolytic
Pol I	0.5	42	ND
Pol II	0.5	100	ND
Pfu	0.5	35.7	ND
Taq	0.5	ND	100

heterotetrameric structure containing two large and two small subunits in one molecule of 421-kDa [27]. Moreover, when either of the two Pol II subunits were individually expressed in *E. coli*, no polymerizing activity could be detected in sonicated or heat-treated crude extracts. However, when the two extracts were mixed together, the activity was restored.

Biochemical properties of *P. abyssi* Pol I and Pol II

We compared substrate specificity, thermostability, pH dependency, salt dependency and sensitivity to aphidicolin of the two DNA polymerases. The pH dependencies of Pol I and Pol II were clearly distinct. Pol I was more active at pH 8.5–9 in a glycine/NaOH buffer, whereas Pol II showed optimal activity in a potassium phosphate buffer at pH 6.5 (Fig. 4A). Tris/HCl buffer seemed to extend the range of activity of both Pol I and Pol II; in 50 mM Tris/HCl buffer, both Pol I and Pol II retained more than 80% of their optimal activity between pH range 7–10 and 7–8.5, respectively (data not shown). We compared the optimal concentrations of $MgCl_2$ and KCl required for the incorporation reaction using activated DNA at the optimal pH. Pol I had an optimum $MgCl_2$ concentration of 3 mM and Pol II had an optimum range of 15–20 mM (Fig. 5A). The optimal concentrations of KCl were 50–80 mM for Pol I and 40 mM for Pol II (Fig. 5B).

The optimal temperature for polymerase activity could not be measured because activated DNA was not stable above 75 °C. The thermostability of Pol I and Pol II was tested at different temperatures: 70, 80, 90 and 100 °C (Fig. 5C). Pol I was found to be more thermotolerant than Pol II.

The sensitivities of the two DNA polymerases to aphidicolin, an inhibitor of many eukaryal-type family B DNA polymerases, were compared. The activity of Pol I was inhibited by aphidicolin (Fig. 4B). However, Pol II was not inhibited at this range of drug concentration. Similar results were obtained with the already described family D DNA polymerases from *P. furiosus* [28] and *Methanococcus jannaschii* [29]. The substrate specificities of Pol I and Pol II are shown in Table 2. Pol I and Pol II utilized various types of DNA as template-primers. The highest activity of *P. abyssi* Pol II was measured with primed ssM13 DNA as a substrate and Pol I preferred the artificial substrate poly(dA)/oligo(dT). In contrast, activity on activated calf thymus DNA was ≈ 2.5-fold lower. These results support the notion that Pol I and Pol II may participate in DNA replication in *P. abyssi*. However, further experiments, specifically designed to study processivity of the two DNA polymerases, will give further insight into their potential cellular function as either a repair or a replicative enzyme. However, it is probable that some other associated proteins function together with these DNA polymerases for the *in vivo* synthesis of DNA.

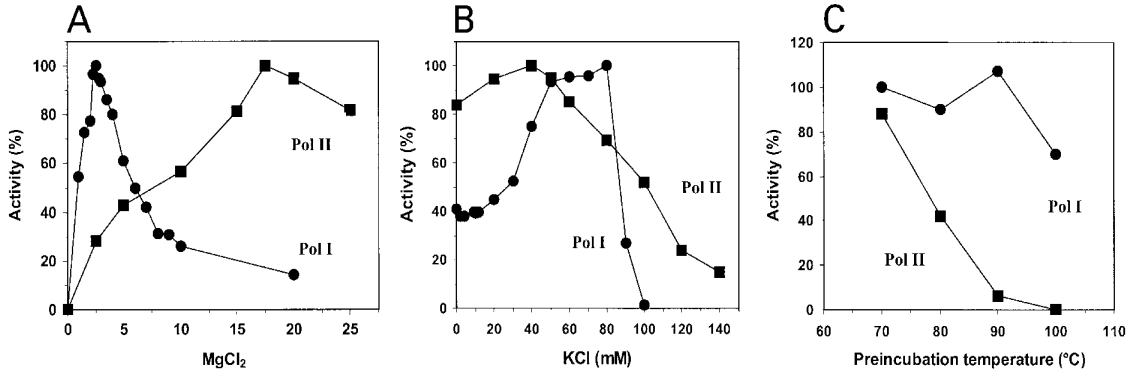

Fig. 5. Influence of magnesium ions (A), potassium ions (B) and thermostability (C) on the activities of *P. abyssi* Pol I and Pol II. (A) The standard assay was carried out at 65 °C for 30 min with various concentrations of $MgCl_2$. One hundred percent activity corresponds to 74 U·mg^{-1} of protein for Pol I and 35 U·mg^{-1} of protein for Pol II. (B) The standard assay was carried out at 65 °C for 30 min with various concentrations of KCl. One hundred percent activity corresponds to 150 U·mg^{-1} of protein for Pol I and 20 U·mg^{-1} of protein for Pol II. (C) The purified DNA polymerase was incubated at various temperatures for 2 h and its activity was measured at 65 °C for 30 min in the standard assay mixture. One hundred percent activity corresponds to 74 U·mg^{-1} of protein for Pol I and 17 U·mg^{-1} of protein for Pol II.

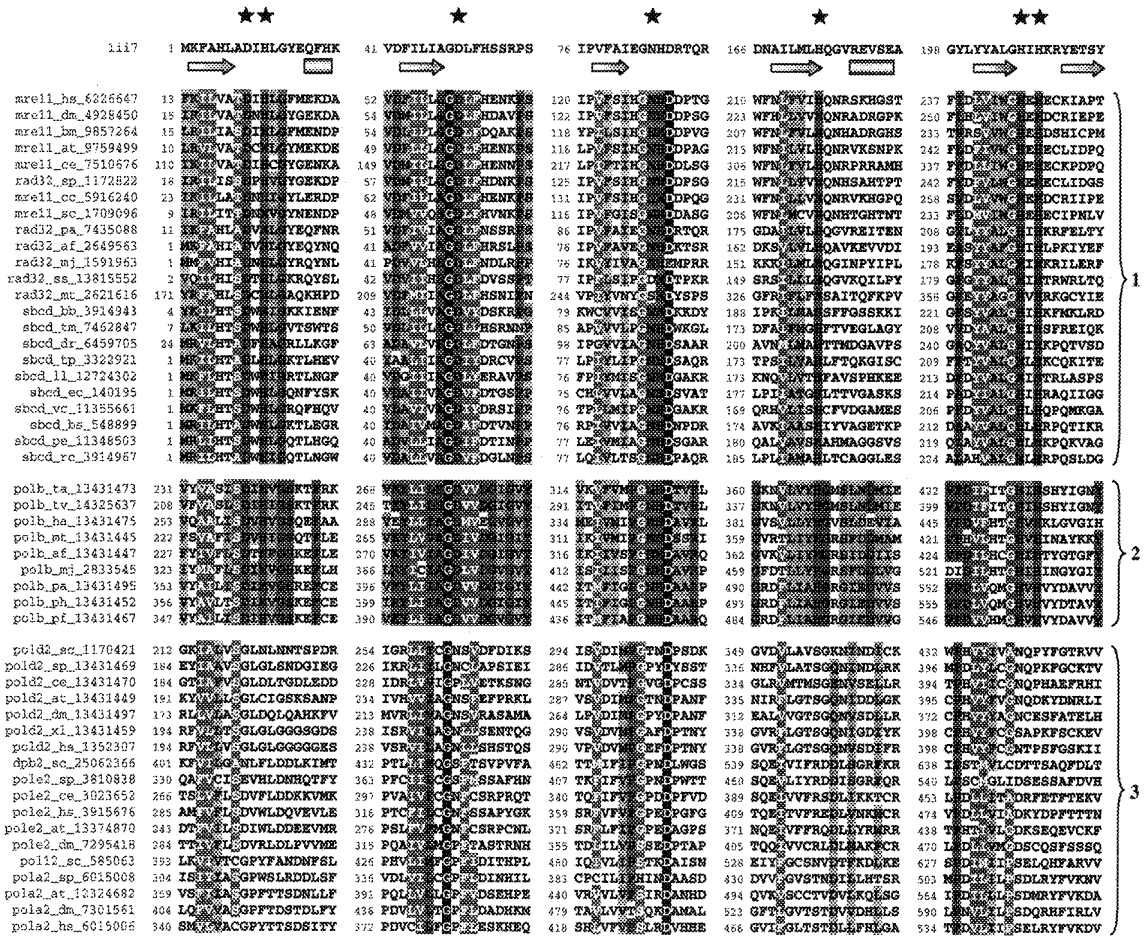

Fig. 6. The five phosphoesterase motifs in nucleases and DNA polymerases. Multiple alignments of the complete sequences have been constructed by DBCLUSTAL [35]. The first line represents the motifs in the Mre11 nuclease from *P. furiosus* for which the X-ray structure has been determined [33] (PDB:1ii7) with the deduced secondary structure (arrow for β strand, rectangle for α helix). The seven residues involved in metal coordination and catalysis are indicated by asterisks. Sequences are divided into three groups: (1) the Mre11/Rad32/Sbcd nuclease family; (2) the small subunit of the archaeal Pol II; (3) the small subunit of eukaryotic polymerases α, δ, ε. In groups 1 and 3, only sequences with < 70% identity are shown because of space limitations. Sequence names are composed of the protein abbreviation, the species abbreviation and the GI number of the original sequence. The numbers before the blocks indicate the position of the first residue in the protein sequence. The shading is based on a 90% consensus. Inverse shading indicates conservation across the three groups (inverse black shading shows residue conservation and inverse grey shading shows physicochemical properties conservation). Grey and light grey shading show residue and physicochemical property conservation within a group, respectively. Physicochemical groups considered are: small (P, G, S, T, A), aromatic (F, Y, W, H), hydrophobic (A, I, L, M, V, F, Y, W), positively charged (K, R, H) and polar residues (D, E, Q, N). The species abbreviations are as follows: hs, *Homo sapiens*; dm, *Drosophila melanogaster*; bm, *Bombyx mori*; at, *Arabidopsis thaliana*; ce, *Caenorhabditis elegans*; sp., *Schizosaccharomyces pombe*; cc, *Coprinus cinereus*; sc, *Saccharomyces cerevisiae*; pa, *P. abyssi*; af, *Archaeoglobus fulgidus*; mj, *Methanococcus jannaschii*; ss, *Sulfolobus solfataricus*; mt, *Methanobacterium thermoautotrophicum*; bb, *Borrelia burgdorferi*; tm, *Thermotoga maritima*; Dr, *Deinococcus radiodurans*; tp, *Treponema pallidum*; ll, *Lactococcus lactis*; ec, *Escherichia coli*; vc, *Vibrio cholerae*; bs, *Bacillus subtilis*; pe, *Pseudomonas aeruginosa*; rc, *Rhodobacter capsulatus*, ta, *Thermoplasma acidophilum*; tv, *Thermoplasma volcanium*; ha, *Halobacterium* sp.; ph, *P. horikoshii*; pf, *P. furiosus*: xl, *Xaenopus laevis*.

Associated exonuclease activities

Many DNA polymerases are known to have associated exonuclease activities. Therefore, we assayed the 5′→3′ and 3′→5′ exonuclease activities of the purified recombinant Pol I and Pol II (Table 3). No 5′→3′ exonuclease activity was detected for Pol I or Pol II. Conversely, both Pol I and Pol II were found to exhibit 3′→5′ exonuclease activity. Pol I and *Pfu* (Promega, used as a control) were shown to have similar exonuclease activity relative to their DNA polymerase activities, whereas the exonuclease activity of Pol II was 2.5-fold stronger (Table 3). Almost all archaeal family B DNA polymerases were known to have associated 3′→5′ exonuclease activity, which is responsible for correction of mismatched dNTPs. Three domains (Exo I, Exo II and Exo III) have been proposed to be essential for this activity [30]. These three domains were identified in the Pol I amino-acid sequence. In contrast, these three domains could not be found within the Pol II amino-acid sequence despite the associated 3′→5′ exonuclease activity. The three described family D DNA polymerases from *P. furiosus* [28], *M. jannaschii* [29] and *P. horikoshii* [27] were also shown to possess associated 3′→5′ exonuclease activity but no domain essential for this activity in this family has been

identified. The large subunit of *P. horikoshii* family D DNA polymerase has been investigated by site-directed mutagenesis of conserved aspartate or glutamate residues but none of the 28 reported mutations provoke the loss of exonuclease activity [27]. Sequence analysis has revealed that the small subunit of the euryarchaeal family D DNA polymerase, as well as the small subunit of eukaryotic pol α, δ and ε, belong to the large calcineurin-like phosphoesterase superfamily defined by five conserved motifs [31]. This superfamily consists of enzymes with a common di-metal active site [32] but with diverse functions such as protein phosphoserine phosphatases, nucleotidases, sphingomielin phosphodiesterases and nucleases. Aravind & Koonin [31] suggested that, in the small subunit of the family D DNA polymerase, the phosphoesterase domain may be involved in pyrophosphate hydrolysis. However, a crystal structure has recently deciphered a di-metal nuclease mechanism in one member of the phosphoesterase superfamily, the Mre11 nuclease of *P. furiosus* [33]. The seven reported residues involved in metal coordination and catalysis are conserved in the small subunit of the family D DNA polymerase but absent in the small subunit of eukaryotic pol α, δ and ε (Fig. 6). An appealing hypothesis is that the phosphoesterase domain of the small subunit is indeed responsible for the 3′→5′ exonuclease activity of archaeal family D DNA polymerase through a similar divalent cation mechanism. In this case, the absence of the seven essential residues in the eukaryotic small subunits would reflect their loss of exonuclease activity. This is in agreement with experimental data: the 3′→5′ exonuclease activity of pol δ and pol ε is provided by the large subunit and no enzymatic activity has been assigned so far to the small subunit [25]. If future site-directed mutagenesis studies confirm our hypothesis, both subunits of the archaeal family D DNA polymerase would be catalytic but with distinct activities: the small subunit carrying the 3′→5′ exonuclease function and and the large subunit carrying the DNA polymerase function.

In vivo function of *P. abyssi* DNA polymerases

The study of DNA replication in general and DNA polymerases in particular in hyperthermophilic archaea is important and interesting in its own right and from an evolutionary point of view. Moreover, the revelation of that the replication machinery of archaea may have many of the basic components of the eukaryal replication machinery (origin recognition, helicase, clamp loader, elongation factor, DNA strand synthesis and ligation, primer synthesis, removal of primers) offers new opportunities for understanding the complexities of eukaryal systems. Studies on DNA replication in archaea have been initiated recently and a few proteins involved in the process have been purified and characterized [24]. The complete *P. abyssi* genome sequence, which allowed the major proteins involved in DNA replication to be identified, combined with the identification of the replication origin [10], should provide important tools for future *in vivo* and *in vitro* studies on DNA replication, repair and recombination in this organism. This report represents the first step in the characterization of the DNA replication process in *P. abyssi*. *P. abyssi* eukaryal homologues of the replication factor C (RFC), proliferating cell nuclear antigen (PCNA) and replication protein A (RPA) have been expressed and characterized in the laboratory. Preliminary results clearly show that, as in eukarya and in the euryarchaeon *P. furiosus* [34], *P. abyssi* PCNA and RFC stimulate DNA synthesis by Pol I and Pol II in *P. abyssi* (J. P. Raffin, unpublished results). In addition, due to the properties of DNA polymerase family D, Cann & Ishino [8] and Shen *et al.* [27] proposed that the family represents the replicative DNA polymerase of euryarchaea. All these recent results suggest that studies on archaea will provide an insight into DNA replication and that further efforts will lead to exciting discoveries.

REFERENCES

1. Jonsson, Z.O. & Hubscher, U. (1997) Proliferating cell nuclear antigen: more than a clamp for DNA polymerases. *Bioessays* **19**, 967–975.
2. Woese, C.R., Kandler, O. & Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl Acad. Sci. USA* **87**, 4576–4579.
3. Kahler, M. & Antranikian, G. (2000) Cloning and characterization of a family B DNA polymerase from the hyperthermophilic crenarchaeon *Pyrobaculum islandicum*. *J. Bacteriol.* **182**, 655–663.
4. De Felice, M., Sensen, C.W., Charlebois, R.L., Rossi, M. & Pisani, F.M. (1999) Two DNA polymerase sliding clamps from the thermophilic archaeon *Sulfolobus solfataricus*. *J. Mol. Biol.* **291**, 47–57.
5. Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D. et al. (1996) Complete genome sequence of the methanogenic Archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073.
6. Klenk, H.P., Clayton, R.A., Tomb, J.F., White, O., Nelson, K.E., Ketchum, K.A., Dodson, R.J., Gwinn, M., Hickey, E.K., Peterson, J.D. et al. (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364–370.
7. Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kogusi, H., Hosoyama, A. et al. (1998) Complete sequence of the gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. *DNA Res.* **5**, 147–155.
8. Cann, I.K.O. & Ishino, Y. (1999) Archaeal DNA replication: Identifying the pieces to solve a puzzle. *Genetics* **152**, 1249–1267.
9. Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A. et al. (1999) Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res.* **6**, 83–101.
10. Myllykallio, H. & Forterre, P. (2000) Mapping of a chromosome replication origin in an archaeon: response. *Trends Microbiol.* **8**, 537–539.
11. Cann, I.K.O., Ishino, S., Nomura, N., Sako, Y. & Ishino, Y. (1999) Two family B DNA polymerases from *Aeropyrum pernix*, an aerobic hyperthermophilic crenarchaeote. *J. Bacteriol.* **181**, 5984–5992.
12. Uemori, T., Ishino, Y., Doi, H. & Kato, I. (1995) The hyperthermophilic archaeon *Pyrodictium occultum* has two alpha-like DNA polymerases. *J. Bacteriol.* **177**, 2164–2177.
13. Edgell, D.R., Klenk, H.P. & Doolittle, W.F. (1997) Gene duplications in evolution of archaeal family B DNA polymerases. *J. Bacteriol.* **179**, 2632–2640.
14. Erauso, G., Reysenbach, A.L., Godfroy, A., Meunier, J.R., Crump, B., Partensky, F., Baross, J.A., Marteinsson, V., Barbier, G., Pace, N.R. et al. (1993) *Pyrococcus abyssi* sp. nov., a new

hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Arch. Microbiol.* **160**, 338–349.
15. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York.
16. Ho, S.N., Pullen, J.K., Horton, R.M., Hunt, H.D. & Pease, L.R. (1990) DNA and protein engineering using the polymerase chain reaction: splicing by overlap extension. *DNA Prot. Eng. Tech.* **2**, 50–55.
17. de Moerlooze, L., Struman, I., Renard, A. & Martial, J.A. (1992) Stabilization of T7-promoter-based pARHS expression vectors using the *parB* locus. *Gene* **21**, 91–93.
18. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
19. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein using the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
20. Henneke, G., Raffin, J.P., Ferrari, E., Jonsson, Z.O., Dietrich, J. & Hubscher, U. (2000) The PCNA from Thermococcus fumicolans functionally interacts with DNA polymerase delta. *Biochem. Biophys. Res. Commun.* **276**, 600–606.
21. Kong, H., Kucera, R.B. & Jack, W.E. (1993) Characterization of a DNA polymerase from the hyperthermophile Archae *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *J. Biol. Chem.* **268**, 1965–1975.
22. Perler, F.B., Davis, E.O., Dean, G.E., Gimble, F.S., Jack, W.E., Neff, N., Noren, C.J., Thorner, J. & Belfort, M. (1994) Protein splicing: inteins and exteins – a definition of terms and recommended nomenclature. *Nucleic Acids Res.* **22**, 1125–1127.
23. Myllykallio, H., Lopez, P., Lopez Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H. & Forterre, P. (2000) Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* **288**, 2212–2215.
24. Kelman, Z. (2000) The replication origin of archaea is finally revealed. *Trends Biochem. Sci.* **25**, 521–523.
25. Makiniemi, M., Pospiech, H., Kilpelainen, S., Jokela, M., Vihinen, M. & Syvaoja, J.E. (1999) A novel family of DNA-polymerase-associated B subunits. *Trends Biochem. Sci.* **24**, 14–16.
26. Cann, I.K.O., Komori, K., Toh, H., Kanai, S. & Ishino, Y. (1998) A heterodimeric DNA polymerase: evidence that members of Euryarchaeota possess a distinct DNA polymerase. *Proc. Natl Acad. Sci. USA* **95**, 14250–14255.
27. Shen, Y., Musti, K., Hiramoto, M., Kikuchi, H., Kawarabayashi, Y. & Matsui, I. (2001) Invariant Asp1122 and Asp1124 are essential residues for polymerization catalysis of family D DNA polymerase from *Pyrococcus horikoshii*. *J. Biol. Chem.* **20**, 27376–27383.
28. Uemori, T., Sato, Y., Kato, I., Doi, H. & Ishino, Y. (1997) A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression and characterization. *Genes Cells* **2**, 499–512.
29. Ishino, Y., Komori, K., Cann, I.K.O. & Koga, Y. (1998) A novel DNA polymerase family found in Archaea. *J. Bacteriol.* **180**, 2232–2236.
30. Blanco, L., Bernad, A., Blasco, M.A. & Salas, M. (1991) A general structure for DNA-dependent DNA polymerases. *Gene* **100**, 27–38.
31. Aravind, L. & Koonin, E.V. (1998) Phosphodiestarase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res.* **26**, 3746–3752.
32. Villafranca, J.E., Kissinger, C.R. & Parge, H.E. (1996) Protein serine/threonine phosphatases. *Cur. Opin. Biotech.* **7**, 397–402.
33. Hopfner, K.P., Karcher, A., Kraig, L., Woo, T.T., Carney, J.P. & Tainer, J.A. (2001) Structural biochemistry an interaction architecture of the DNA double-stand break repair Mre11 nuclease and Rad50-ATPase. *Cell.* **105**, 473–485.
34. Cann, I.K.O., Ishino, S., Hayashi, I., Komori, K., Toh, H., Morikawa, K. & Ishino, Y. (1999) Functional interactions of a homolog of proliferating cell nuclear antigen with DNA polymerases in Archaea. *J. Bacteriol.* **181**, 6591–6599.
35. Thompson, J.D., Plewniak, F., Thierry, J. & Poch, O. (2000) DbClustal: rapid and reliable global multiple alignments of protein sequences detected by database searches. *Nucleic Acids Res.* **28**, 2919–2926.